Predicting the outcome of a game

David H. Wolpert[a]

[a]NASA Ames Research Center, MS 269-1, Moffett Field, CA, 94035
dhw@ptolemy.arc.nasa.gov

Summary. The positive interpretation of conventional game theory predicts that the joint (mixed) strategy of players in a game satisfies an equilibrium concept. The relative probabilities of such satisfying strategies are not specified, and all other strategies are deemed impossible. As an alternative, in this paper we use statistical inference to predict the joint strategy. The associated positive problem is to determine the density function over joint strategies.

This Predictive Game Theory (PGT) typically assigns non-zero probability density to multiple joint strategies. A loss function can be used to distill that density to a single joint strategy via decision theory. This mapping of a game to a joint strategy constitutes an "equilibrium concept". It typically produces a single joint strategy and therefore needs no refinements.

We explore a Bayesian version of PGT based on the entropic prior and a likelihood that quantifies the rationalities of the players. We show that the local peaks of the posterior density and the game's Quantal Response Equilibria (QRE's) approximate each other. Some joint strategies are not expressible as QRE's. In contrast, we show that every joint strategy has non-zero density for one (and only one) set of player rationalities.

Keywords: Quantal Response Equilibrium, Bayesian Statistics, Entropic prior, Statistical Physics

JEL Classification codes: C72

1 Introduction

Say one wishes to predict some characteristic of interest y concerning some physical system. To make the prediction one is provided some information $\mathscr{I}$ concerning the system. Statistical analysis starts by converting $\mathscr{I}$ into a probability distribution over y.[1] Such a distribution is far more informative than a single "best prediction". However if needed we can synopsize the distribution with a single prediction. One way to do that is to use the mode of the distribution as the prediction. When the distribution is a Bayesian posterior probability, $P(q \mid \mathscr{I})$, this mode is called the Maximum A Posterior (MAP) prediction. Alternatively, say there is a real-valued loss function, $L(y, y')$ that quantifies the penalty we will incur if we predict y' and the true value is y. Then Bayesian decision theory counsels us to predict the value of y' that minimizes the posterior expected loss, $\int dy L(y, y') P(y \mid \mathscr{I})$ [Jaynes and Bretthorst, 2003, Gull, 1988, Loredo, 1990, Bernardo and Smith, 2000, Berger, 1985, Zellner, 2004, Paris, 1994, Horn, 2003].

This standard approach to analyzing physical systems can be appropriate even when the system being analyzed is a set of human beings playing a game. In particular, one can apply this approach to non-cooperative strategic games, by taking the "characteristic of interest" y to be the joint mixed strategy, q. In this application the Bayesian posterior is a distribution over joint mixed strategies, $P(q \mid \mathscr{I})$.

We use the term **Predictive Game Theory** (PGT) to refer to any such application of statistical inference *to* games (in contrast to the use of statistical inference by some players *in* a game). In this paper we focus on PGT for non-cooperative strategic form games. For simplicity we will adopt the Bayesian approach to inferring a distribution over q's, although non-Bayesian approaches could also be used.

The posterior assigns probability values to all joint strategies q. This contrasts to what conventional equilibrium concepts provide, which is a subset of all possible q's with no associated probability values (except in the degenerate sense that if that set contains a single element we ca interpret it as having probability 1.0). Due to this difference, PGT allows more sophisticated tests comparing experiment and theory than do conventional equilibrium concepts, e.g., tests of theoretical predictions concerning the variances of various attributes of the players' behavior.

Nonetheless, in practice sometimes one must produce a single joint strategy as one's "prediction" of the joint strategy. To that end, say that we have a loss function $L(q', q)$ that quantifies the penalty we will incur if we predict the joint mixed strategy q' and the true joint mixed strategy is q. Then decision theory counsels us to set our single prediction to the "Bayes-optimal" joint mixed strategy, i.e., to the q' that minimizes expected $L(q', q)$ under the posterior density over q. By mapping a game to a single predicted joint strategy this way, decision theoretic PGT provides an "equilibrium concept". Unlike typical equilibrium concepts [Fudenberg and Tirole, 1991, Aumann and Hart, 1992,

[1] We will sometimes be loose in distinguishing between probability distributions, probability density functions, etc., and generically we will write "$P(\ldots)$" to mean whichever is appropriate.

Basar and Olsder, 1999, Binmore, 1992, Luce and Raiffa, 1985], this one does not require refinements; typically the Bayes-optimal prediction is unique.

The Bayes-optimal equilibrium concept depends on the loss function of the external statistician. An alternative "equilibrium concept" that does not depend on a loss function is given by marginalizing the posterior over densities $q(x)$, $P(q \mid \mathscr{I})$, down to a (single) posterior over joint pure strategies $x = (x_1, x_2, \ldots)$ (joint moves), $P(x \mid \mathscr{I})$. Typically the moves of the players are stochastically coupled under that posterior over joint moves. Moreover, often the marginal distribution over a particular player's moves, x_i, is not utility-maximizing against the marginal distribution over the other players' moves, This can occur even if the support of the posterior over joint mixed strategies is restricted to Nash equilibria. In this sense, Nash equilibria may be impossible, and bounded rationality is unavoidable.

In this paper for pedagogical simplicity we investigate the Bayesian posterior for an entropic prior over mixed strategies. For similar reasons we pick a likelihood that says, in essence, that the logit Quantal Response Equilibria (QRE's [J. K. Goeree, 1999, McKelvey and Palfrey, 1995, Chen and Friedman, 1997]) are, at the very least, not inconsistent with human behavior in games against Nature. With these choices the QRE's turn out to be the local peaks of the posterior for games against Nature. In particular the MAP distribution is a QRE for such a game. However with the same choice for the likelihood, the QRE's are only approximations to the local peaks of the posterior for strategic games involving other goal-seeking players.

Our choice for the likelihood means that it is parameterized by the usual QRE parameters (the exponents in the logit distributions of the players), which we will call the "rationalities" of the players. Some joint mixed strategies are not a QRE for any choice of the player rationalities. For example, this can be the case for certain Nash equilibria. However we show here that every joint strategy — including every Nash equilibrium — has non-zero posterior for an appropriate set of rationalities of the players.

In the remainder of this introduction we present some notation. In the following section we discuss general aspects of PGT and the QRE. In particular we discuss their relation to bounded rationality. In the section after that we review the entropic prior and some elementary properties of logit distributions. In the next section we present our likelihood function based on the QRE for a game against Nature. We then discuss the posterior over joint mixed strategies given by combining that likelihood with the entropic prior. In particular, we derive sufficient conditions for the QRE's of an N-player game to be the MAP's of the posterior over joint mixed strategies. In our final section we briefly survey some applications of PGT not presented in this paper.

1.1 Notation

We consider a general noncooperative game that has N independent **players**, indicated by the natural numbers $\{1, 2, \ldots, N\}$. Each player i has the finite set of allowed pure strategies $x_i \in X_i$, where $|X_i|$ is the (finite) cardinality of X_i. The set of all possible joint strategies is $X \triangleq X_1 \times X_2 \times \ldots \times X_N$ with cardinality $|X| \triangleq \prod_{i=1}^{N} |X_i|$, a generic element of X being written as x. $u^i : X \to \mathbb{R}$ is player i's utility function, the mixed strategy of i is the distribution $q_i(x_i)$, and $q(x) \triangleq \prod_{i=1}^{N} q_i(x_i)$.

$\Delta_{\mathcal{X}}$ is the Cartesian product of the simplices Δ_{X_i} (implicitly imbued with the standard product topology over simplicial complexes). So mixed joint strategies (i.e., product densities) are elements of $\Delta_{\mathcal{X}}$. The expected utility of player i is written as $\mathbb{E}_q(u^i) = \sum_x \prod_j q_j(x_j) u^i(x)$. We define each player i's **environment** function, often with the associated random variable q_{-i} implicit, as $U^i_{q_{-i}}(x_i) \triangleq \mathbb{E}_{q_{-i}}(u^i \mid x_i)$. We will sometimes write $\mathbb{E}_q(u^i) = q_i \cdot U^i$.

Cov is the covariance operator, defined for any countable set of variables $\{y\}$ and associated distribution $p \in \Delta_Y$ by

$$\mathrm{Cov}_p[a(y), b(y)] \triangleq \sum_{y \in Y} p(y)a(y)b(y) \; - \; \sum_y p(y)a(y) \sum_y p(y)b(y).$$

(For clarity, we will sometimes write this as $\mathrm{Cov}_{p(y)}[a(y), -ib(y)]$.) Given any player i, we will use $-i$ to refer to the set of all $N-1$ other players. In particular, we will sometimes write $q_{-i} \times q_i$ to indicate the $p \in \Delta_{\mathcal{X}}$ with components $p(x) = p(x_i, x_{-i}) \triangleq q_i(x_i)q_{-i}(x_{-i})$.

Curly braces indicate an entire set and vertical bars the cardinality of a finite set, e.g., $\{\beta_i\}$ is the set of all values of β_i for all i, and $|\{\beta_i\}|$ the number of such i. Bold letters, e.g., $\boldsymbol{a}$, mean a finite-dimensional vector over the extended real numbers $\mathbb{R}^*$ (i.e., the reals together with positive and negative infinity [Aliprantis and Border, 2006]). $\boldsymbol{a} \succeq \boldsymbol{b}$ indicates the generalized inequality that $\forall i$, $a_i \geq b_i$. $I(.)$ is the indicator function that equals 1 if the equation that is its argument is true and 0 otherwise.

Just as "$P(.)$" means a distribution or density function as appropriate, so "$\delta(.)$" indicates the Dirac or Kronecker delta function, as appropriate. Also, we will sometimes refer to "$\mathbb{R}^+$", "$[1.0, \infty)$", etc., when what we really mean are, respectively, the non-negative extended reals, the extended reals greater than or equal to 1.0 (including $+\infty$), etc.

To distinguish it from densities like q, a distribution (density function) P that describes our prediction is called a **predictive distribution** (density function).[2] So for example, $P(q \mid \mathscr{I})$, $P(x \mid \mathscr{I}) = \int dq \; P(q \mid \mathscr{I})q(x)$, and $P(q \mid \mathscr{I}, x_j)$ are all predictive densities. Predictive densities reflect *our* knowledge/insight/ignorance concerning the game. This contrasts with distributions like q, which reflect the "physical" distributions of the players in the game.

[2] This use of the term "predictive distribution" should not be confused with the one arising in Bayesian statistics.

2 Equilibrium concepts and bounded rationality

2.1 The two equilibrium concepts of PGT

Say we have information $\mathscr{I}$ about a game involving a set of human players. We want to predict what mixed joint strategy q those humans will play. Adopting the role of a Bayesian statistician external to the physical system of those humans, to make this prediction means determining $P(q \mid \mathscr{I})$. As described in the introduction, if we have a loss function, then decision theory provides us an "equilibrium concept" for the game by mapping $P(q \mid \mathscr{I})$ to an associated Bayes-optimal prediction for the joint mixed strategy q.

The introduction also mentioned a second way that the posterior over $\Delta_{\mathcal{X}}$ induces a single distribution over X. This second "equilibrium concept" is

$$\begin{aligned} P(x \mid \mathscr{I}) &= \int dq \; P(x \mid q, \mathscr{I}) P(q \mid \mathscr{I}) \\ &= \int dq \; q(x) P(q \mid \mathscr{I}). \end{aligned} \tag{1}$$

Both equilibrium concepts reflect two kinds of ignorance. The first kind of ignorance is that of us, the external statistician, concerning the game and its players. This kind of ignorance is encapsulated in P. The second kind is the intrinsic randomness in how the players choose their moves, and is encapsulated in each q.

The decision-theoretic equilibrium concept varies with the loss function, unlike the $P(x \mid \mathscr{I})$ equilibrium concept. However while a decision-theoretic equilibrium is always a product distribution, $P(x \mid \mathscr{I})$ may not be, i.e., X_i and X_j may be statistically dependent given only $\mathscr{I}$. This is true even though the support of $P(q \mid \mathscr{I})$ is restricted to distributions where X_i and X_j are independent (a linear combination of product distributions typically is not a product distribution). In addition, say that $P(q \mid \mathscr{I})$ is restricted to Nash equilibria q. Typically, if there are multiple such equilibria, then $P(x_i \mid \mathscr{I})$ is not an optimal response to $P(x_{-i} \mid \mathscr{I})$. Even if we know that all the players are perfectly rational, *our prediction* of their moves has "cross-talk" among the multiple equilibria, which prevents perfect rationality. This is one sense in which PGT has built-in bounded rationality.

Example 1: To illustrate the foregoing, consider a two player game in which both players have two possible moves, L and R. Indicate any (product distribution) q by two numbers, $q_1(x_1 = L)$ and $q_2(x_2 = L)$. Suppose that

$$P(q \mid \mathscr{I}) = \frac{\delta(q - (3/4, 3/4)) + \delta(q - (1/4, 1/4))}{2} \tag{2}$$

where "$\delta(.)$" is the Dirac delta function. Suppose also that we have quadratic loss. For that loss function, as is easy to verify, the Bayes-optimal q is the average q, $\int dq\; q(x)P(q \mid \mathscr{I})$. Viewed as a function of x, that particular Bayes-optimal q is the same as $P(x \mid \mathscr{I})$. Here they equal the distribution $P(L, L) = P(R, R) = 5/16, P(R, L) = P(L, R) = 3/16$. Indicate that distribution as p. p is not a product distribution, so $P(p \mid \mathscr{I}) = 0$. In other words, $P(x \mid \mathscr{I})$, this game's "equilibrium", is a joint mixed strategy that cannot arise.

Suppose that our information $\mathscr{I}$ concerning a game does not *explicitly* tell us that the players in the game are all fully rational. Then the rationalities of the (human) players are random variables, and we must average over them to get the posterior over joint mixed strategies. This generically means that $P(q \mid \mathscr{I})$ is non-zero for joint mixed strategies q that are not perfectly rational. This is another way that PGT provides built-in bounded rationality.

To help distinguish when one should use one equilibrium concept or the other, consider a frequentist scenario, where we first give our prediction $q' \in \Delta_{\mathcal{X}}$ for the outcome of a game, and after that $P(q \mid \mathscr{I})$ is IID sampled an infinite number of times. If our reward for making prediction p is the average value of $L(q', q)$ over that infinite number of samples, then to maximize our reward we should use this Bayes-optimality equilibrium concept.

Say that instead, each time $P(q \mid \mathscr{I})$ is sampled to produce a q, that that q is itself sampled, to produce an x. This means that the IID samples of $P(q \mid \mathscr{I})$ provide an empirical distribution of the frequency with which each x occurs. With probability 1.0, the uniform metric distance between this empirical distribution and $P(x \mid \mathscr{I}) = \int dq\; P(q \mid \mathscr{I})q(x)$ is zero. But that integral is just the second equilibrium concept discussed above. So if the reward is how accurately we guess the empirical distribution over x's, then we should use this second equilibrium concept instead of the Bayes-optimality equilibrium concept.

2.2 PGT as a meta-game

The first type of equilibrium concept in PGT can also be motivated as a "meta-game" played against Nature. To formalize this meta-game, say we have a set of possible games G, differing in their utility functions, their players, etc. For each such game $\gamma \in G$, let $\Delta_{\mathcal{X}}(\gamma)$ indicate the set of all possible joint mixed strategies in γ, with $x \in X(\gamma)$ the possible joint moves in that game. Now consider a two-stage "meta-game" Γ that consists of a statistician (S) playing against Nature (N). In this meta-game N's set of possible moves is $\{(\gamma \in G, q \in \Delta_{\mathcal{X}}(\gamma))\}$, i.e., the set of all possible games γ, and for each such game, the set of all possible joint mixed strategies q over the joint moves in γ. The mixed strategy of player N is a distribution over this space, $P(\gamma \in G, q \in \Delta_{\mathcal{X}}(\gamma))$. At the end of the first stage of the meta-game, N's mixed strategy is fairly sampled, producing an outcome (γ', q').

As an example, if the players in the underlying game γ' are perfectly rational, then the support of $P(q' \mid \gamma')$ is the Nash equilibria of γ'. *A priori*, there is no problem of deciding among those equilibria (the traditional motivation for equilibrium refinements). All of them can occur, with relative probabilities given by $P(q' \mid \gamma')$.

However in the real world sometimes a statistician must produce a single prediction rather than a full posterior density over predictions. This provides a second stage for our meta-game. In this second stage S is told both γ' and N's mixed strategy, $P(\gamma \in G, q \in \Delta_{\mathcal{X}}(\gamma))$. Together, those give S a posterior over what q' is, $P(q' \mid \gamma')$. In the second stage, S makes a move in $\Delta_{\mathcal{X}}(\gamma')$, i.e., picks a joint mixed strategy for the game γ'. We interpret that move of the statistician S as a prediction of what $q' \in \Delta_{\mathcal{X}}(\gamma')$ was produced by the sampling of player N.

As usual in games against Nature, N has no utility function. However S may have a utility function, given by the negative of a loss function $L(q, q')$ that quantifies how accurate her move q is as a prediction of N's move q'. In this case, to maximize her expected utility the statistician chooses her move — her prediction of the joint mixed strategy that governs the game γ' — to minimize her expected loss under the posterior $P(q' \mid \gamma')$. Formally, she should guess a distribution whose support is $\text{argmin}_p \mathbb{E}_{P(q'|\gamma')} L(p, q')$. This mapping of a game γ' to a single predicted joint strategy comprises the first equilibrium concept described above.

2.3 The QRE

There is a rich history of work on encapsulating bounded rationality without recourse to posterior distributions over joint mixed strategies, e.g., stochastic preference theory, non-expected utility theory, behavioral game theory in general and prospect theory in particular, etc. (See [Starmer, 2000, Camerer, 2003, Kahneman, 2003, Kurzban and Houser, 2005, Fudenberg and Levine, 1998, List and Haigh, 2005] for excellent overviews of this work.) Of particular interest in this paper is the approach to bounded rationality embodied in the Quantal Response Equilibrium (QRE). This approach is a modification of a conventional equilibrium concept where one simultaneously models every players i in a game as playing a mixed strategy $q_i(x_i)$ that is a logit (Boltzmann) distribution in her expected utilities. More precisely, one predicts that the outcome of the game is a solution to the simultaneous set of equations

$$q_i(x_i) \propto e^{\beta_i \mathbb{E}_q(u^i | x_i)} \; \forall i \tag{3}$$

where the joint distribution $q(x) = \prod_i q_i(x_i)$.

Not all q can be cast as a QRE for some appropriate $\{\beta_i\}$ (see Sec. 4.2 below). So in particular, a q that occurs in the real world will in general differ, even if only slightly, from all possible QRE's. This can be viewed as a shortcoming of the QRE (a shortcoming of all equilibrium concepts with a small number of parameters).

We use the notation that $q^*_{\{\beta_i\}}(x)$ means a QRE, where the parameters $\{\beta_i\}$ are often implicit. In general, for any particular game and (non-negative) $\{\beta_i\}$, there is at least one (and may be more than one) associated q^*.

This follows from Brouwer's fixed point theorem [McKelvey and Palfrey, 1995, Wolpert, 2004a].

At a Nash equilibrium each player i sets her strategy q_i to maximize her expected utility $\mathbb{E}_{q_i,q_{-i}}(u^i) = \mathbb{E}_{q_i}(U^i_{q_{-i}}$ for fixed q_{-i}. Consider instead having each player i set q_i to maximize an associated functional, the **free utility**:

$$\mathscr{F}_{U^i_{q_{-i}},T_i}(q_i) \triangleq \mathbb{E}_{q_i}(U^i_{q_{-i}}) + T_i S(q_i). \tag{4}$$

For all $T_i \to 0$ the equilibrium q that simultaneously minimizes $\mathscr{F}_i \; \forall i$ is a Nash equilibrium [Wolpert, 2004a, McKelvey and Palfrey, 1995, Meginniss, 1976, Fudenberg and Kreps, 1993, Fudenberg and Levine, 1993, Luce, 1959]. For $T_i > 0$ one gets bounded rationality. Indeed, under the identity $T_i \triangleq \beta_i^{-1} \; \forall i$ the solution to this modified Nash equilibrium concept is a QRE.[3]

In the context of game theory, the free utility Lagrangian was investigated in [Fudenberg and Kreps, 1993, Fudenberg and Levine, 1993, Shamma and Arslan, 2004]. The first attempt to derive it in the game theory context from first principles was in [Meginniss, 1976].

Historically, the QRE was not motivated in terms of free utilities but by modeling payoff uncertainty [McKelvey and Palfrey, 1995]. It can also be motivated as the equilibrium of a learning process by the players, a process that is closely related to replicator dynamics [Wolpert, 2004b, Anderson et al., 2002, Goeree and Holt, 1999]. In addition, in a non-game-theory context, the QRE can be derived from first principles as a way to do distributed control [Wolpert and Rajnarayan, 2007, Wolpert et al., 2006].

Finally, there has been a large body of work relating economics and statistical physics [Brock and Durlauf, 2001, Durlauf, 1999, Dragulescu and Yakovenko, 2000, Aoki, 2004, Farmer et al.,]. (Indeed, there is now an entire field of "econophysics".) Since the logit distribution is the cornerstone of statistical physics (where it occurs in the "canonical ensemble" and the "grand canonical ensemble"), the QRE is also connected to statistical physics. In particular, consider a team game (all u^i are the same) with all players sharing the same rationality. As discussed in [Wolpert, 2004a, Wolpert, 2005], for such a game the (shared) free utility essentially becomes what in statistical physics is known as a "mean field approximation" to the "free energy" of a system (hence the terminology).

3 Mathematical review

Before investigating the relationship between the QRE and PGT we need to review some elementary mathematical tools.

[3] In [McKelvey and Palfrey, 1995], $U^i_{q_{-i}}$ is called "a statistical reaction function", and the set of coupled equations giving that solution is called the "logit equilibrium correspondence".

3.1 Review of the entropic prior

Shannon was the first person to realize that based on any of several separate sets of very simple desiderata, there is a unique real-valued quantification of the amount of syntactic information in a distribution $P(y)$. He showed that this amount of information is (the negative of) the *Shannon entropy* of that distribution, $S(P) = -\sum_y P(y)\ln[\frac{P(y)}{\mu(y)}]$.[4] Note that for a product distribution $P(y) = \prod_i P_i(y_i)$, entropy is additive: $S(P) = \sum_i S(P_i)$. So for example, the distribution with minimal information is the one that doesn't distinguish at all between the various y, i.e., the uniform distribution. Conversely, the most informative distribution is the one that specifies a single possible y.

Say that the possible values of the underlying variable y in some particular probabilistic inference problem have no known *a priori* stochastic relationship with one another. For example, y may not be numeric, but rather consist of the three symbolic values, {red, dog, Republican}. Then simple desiderata-based counting arguments can be used to conclude that the prior probability of any distribution $p(y)$ is proportional to the **entropic prior**, $\exp(\alpha S(p))$, for some associated finite constant $\alpha \geq 0$. [5]

Intuitively, this prior says that absent any other information concerning a particular distribution p, then the larger its entropy the more *a priori* likely it is. Independent of the entropic prior's desideratum-based motivations, it has proven has been very successful in practice [Mackay, 2003, Gull, 1988]. Indeed, it can be used to derive statistical physics, whose predictions have been exhaustively tested [Jaynes, 1957].

Under the entropic prior the posterior probability of p given information $\mathscr{I}$ concerning p is

$$P(p \mid \mathscr{I}) \propto \exp(\alpha S(p)) P(\mathscr{I} \mid p). \tag{5}$$

The associated MAP prediction of p is $\text{argmax}_p P(p \mid \mathscr{I})$. As an example, say that $\mathscr{I}$ is a particular element of a partition on the space of possible p's, i.e., a restriction of p to some particular set. Then for any $\alpha > 0$, the MAP p is the one that maximizes $S(p)$, subject to being one of the p's delineated by $\mathscr{I}$.

Intuitively, Eq. 5 pushes us to be conservative in our inference. Of all hypotheses p equally consistent (probabilistically) with our provided information, we are led to view as more *a priori* likely those p that contain minimal

---

[4] μ is an *a priori* measure over y, often interpreted as a prior probability distribution. Unless explicitly stated otherwise, here we will always assume it is uniform, and not write it explicitly. See [Jaynes, 1957, Jaynes and Bretthorst, 2003, Cover and Thomas, 1991].

[5] The issue of how to choose α for a particular application — or better yet integrate over it — is subtle, with a long history. See work on ML-II [Berger, 1985] and the "evidence procedure" [Strauss et al., 1994].

extra information beyond that provided in $\mathscr{I}$.[6] For this reason, the entropic prior has been proposed as a formalization of Occam's razor.

Note that the entropic prior evaluated for a product distribution is itself a product, i.e., if $q(x) = \prod_i q_i(x_i)$, then $e^{\alpha S(q)} = \prod_i e^{\alpha S(q_i)}$. As a result, by symmetry the associated marginal over x,

$$\sum_x q(x)P(q) \propto \sum_x \prod_i q_i(x_i) e^{\alpha S(q_i)}, \tag{6}$$

must be uniform over x.

3.2 Miscellaneous properties of logit distributions

Certain simple identities and associated definitions concerning logit distributions will prove useful below. First, given any function $f : Y \to \mathbb{R}$ and $c \in \mathbb{R}$, as in statistical physics we define the associated **partition function**

$$Z_f(c) \triangleq \sum_y e^{cf(y)} \tag{7}$$

where we implicitly assume that f is bounded. For finite c, the logit (Boltzmann) distribution in values of $f(y)$ having exponent c is defined by

$$\mathcal{L}_{f,c}(y) \triangleq e^{cf(y)}/Z_f(c) \tag{8}$$

for finite c, and for infinite c by $\mathcal{L}_{f,\infty}(y) \triangleq \delta(y, \text{argmax} f(.))$, $\mathcal{L}_{f,-\infty}(y) \triangleq \delta(y, \text{argmin} f(.))$. Note that for any c and f, $\mathcal{L}_{f,c}(y)$ is uniform over any set of y sharing the same value for $f(y)$. We define the **Boltzmann utility** as

$$K(f,c) \triangleq \sum_y f(y)\mathcal{L}_{f,c}(y). \tag{9}$$

$K(f,c)$ is the expected value of f under the logit distribution in values of f having (potentially infinite) exponent c. The function $K(f,.) : \mathbb{R} \to \mathbb{R}$ is C^∞. Moreover, for any $c \in \mathbb{R}^*$, $K(.,c) : \mathbb{R}^{|Y|} \to \mathbb{R}$ is continuous. (It can be nondifferentiable for infinite c at the point where $f(y) = f(y')$ for some pair $(y, y' \neq y)$.)

A crucial identity in statistical physics which we will also use here gives the first moment of f under the logit distribution:

$$K(f,c) = \frac{d\ln[Z_f(c)]}{dc}. \tag{10}$$

[6] This is different from saying that the larger s is, the more *a priori* likely it is that the system has that s: $P_S(s) = \int dp \; \delta(S(p) - s)P(p) = \frac{\int dp \; \delta(S(p)-s)\exp\,(\alpha S(p))}{\int dp \; \exp\,(\alpha S(p))}$ which may actually decrease with increasing s, depending on the nature of $\frac{dS}{dp}$.

Similarly, the variance of f under the logit distribution over $f(y)$ values equals the second derivative of $\ln[Z_f(c)]$ with respect to c. This variance is strictly positive for finite c and non-constant f. So for such f, $K(f,.)$ is a nowhere decreasing bijection from $\mathbb{R}^* \to [\min_y f(y)), \max_y f(y))]$.

We will often want to find the $p \in \Delta_Y$ that maximizes $S(p)$ subject to the constraint that $\sum_y f(y)p(y) = k$. The (unique) solution is the logit distribution $\mathcal{L}_{f,c}$ where c is a Lagrange parameter set to enforce the constraint, i.e., set so that $K(f,c) = k$. For example, consider maximizing the entropy of a player with distribution p in a game against Nature subject to a provided expected value of that player's utility function, $k = \mathbb{E}_p(f)$. The Lagrangian for this problem is the free utility of the player, $\mathscr{F}_{f,c}(p)$. As mentioned at the end of Sec. 2.3, the associated solution for p is the QRE. In this game-against-Nature context, that is just the logit distribution $\mathcal{L}_{f,c}$.

Since for the proper value of c, $\mathscr{L}_{f,c}$ is the maximizer over $p \in \Delta_Y$ of $S(p)$ subject to the constraint $f \cdot p = k$, it is also the maximizer of $S(p)$ subject to the constraint $f \cdot p = K(f,c)$. This can be used to show that the entropy of the logit distribution $\mathcal{L}_{f,c}$ cannot increase as c rises.[7] So the picture that emerges is that as c increases, the logit distribution gets more peaked, with lower entropy. At the same time, it also gets higher associated expected value of f.

4 The posterior for the entropic prior

The posterior over possible joint mixed strategies q is given by the prior and the likelihood. For pedagogical simplicity we have adopted the entropic prior. This means that if we know nothing about the players in a game (so in particular we do not know their utility functions, their rationalities, etc.), then we view a particular almost uniform joint mixed strategy q as *a priori* more likely than a particular highly peaked joint mixed strategy q. α quantifies how much more likely we find such relatively uniform q.

Given a choice for the prior, our next task is to choose the likelihood, i.e., to formalize what we know about the human players *a priori*.

[7] To see this say we replace the invariant $p \cdot f = K(f,c)$ with $p \cdot f \geq K(f,c)$. Entropy is a concave function of its argument, as is this inequality constraint, so our new optimization problem is concave. Therefore the critical point of the associated Lagrangian is the optimizing p. Now if we increase c, and therefore increase $K(f,c)$, the feasible region for our new invariant decreases. This means that when we do that the maximal feasible value of S cannot increase. So the entropy of the critical point of the Lagrangian for our new invariant cannot increase as c does. However that critical point is just the logit distribution $p = \mathcal{L}_{f,c}$, i.e., it is the optimizing p for the original equality invariant, $p \cdot f = K(f,c)$. So the property that increasing c cannot increase the entropy under the new invariant must also hold for the original equality invariant.

4.1 The likelihood

The first thing we know about the players is that under their joint mixed strategy their moves are statistically independent (since we are restricting attention to normal form games). Beyond that, all of the insights of behavioral game theory, psychology, and human modeling [Camerer, 2003, Starmer, 2000, Allais, 1953, List and Haigh, 2005, Kurzban and Houser, 2005] could be brought to bear on the task of determining the likelihood.

Here though we will not try to formalize those insights. Instead we will simply assume that any player will try to maximize her expected utility, to the best of her computational abilities, the best of her insights into the other players and the game structure, etc. To formalize this minimal assumption, first consider just those instances in which player i is confronted with some single environment $U^i_{q_{-i}}$. We assume that on average, the move i chooses results in the same utility in all those instances: $q_i \cdot U^i_{q_{-i}}$ has the same (potentially unknown) value in all of them. We write that value as $\epsilon_i(U^i_{q_{-i}})$. $\mathscr{I}$ is the restriction that q is a product distribution and that simultaneously for all players i, $q_i \cdot U^i_{q_{-i}} = \epsilon_i(U^i_{q_{-i}})$. As an example, at a Nash equilibrium $\epsilon_i(U^i_{q_{-i}}) = \max_{x_i} U^i_{q_{-i}}(x_i) \; \forall i$.

This likelihood amounts to saying that as far as player i is concerned when she chooses her move, there is only one salient aspect of q_{-i}. That salient aspect is the effect of q_{-i} on the utility values for i's possible moves, i.e., its effect on $U^i_{q_{-i}}$. The likelihood embodies this aspect of q_{-i} and ignores all other (non-salient) aspects of q_{-i}. In stipulating that only the effects of q_{-i} on her utility are salient to any player i, this likelihood follows the spirit of the axioms of utility theory.

Our next step is to specify the function ϵ_i. To do this we consider how player i would behave in a counterfactual "game against Nature". In that new problem we focus on just one player i, fixing the mixed strategies of the others, so that there are no common knowledge issues, no reasoning about the reasoning of others. Our presumption is simply that any player i's expected utility in such a game against Nature is consistent with what it would be if — as in a QRE — she were using a logit mixed strategy for some associated exponent b_i. In essence, we assume that at the very least, the QRE is consistent with a player's expected utility in games against Nature, i.e., that its likelihood is non-zero in such a game.

We refer to the exponent b_i in i's likelihood as i's **rationality**. There is only one QRE for a game against Nature, namely $q^*_i = \mathcal{L}_{U^i_{q_{-i}}, b_i}$. Since ϵ_i must give the expected value of $U^i_{q_{-i}}$ under this distribution,

$$\epsilon_i(f_i) \triangleq K(f_i, b_i) \tag{11}$$

for some appropriate constant b_i. Our likelihood for this game against Nature is that (q is a product distribution and that) $q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$ with q_{-i}

being fixed. So the posterior is

$$P(q \mid \mathscr{I}) \;\propto\; e^{\alpha S(q)} I(q_i \cdot U_{q_{-i}} = K(U_{q_{-i}}, b_i)) \prod_{j \neq i} \delta(q_j - q'_j) \tag{12}$$

where $\{q'_j : j \neq i\}$ are the pre-fixed distributions of all players other than i. The MAP for this game against Nature equals the associated QRE:

Proposition 1. *For a game of player i against Nature, there is a single local peak of the posterior over q_i for rationality b_i, and there is only one QRE q_i for logit exponent b_i, and those two q_i's are identical.*

Proof. For $\alpha > 0$ any local peak of the posterior is a distribution $q = (q_i, q'_{-i})$ that maximizes $S(q)$ subject to the constraint that $q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$. Since S is additive for product distributions, this q is given by the q_i's that maximize $S(q_i)$ subject to the constraint that $q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$. As described at the end of Sec. 3.2, there is a unique such local peak q_i, given by the logit distribution $\mathcal{L}_{\mathbb{E}_{q_{-i}}(u^i|.), b_i}(x_i)$. This proves the claim for $\alpha > 0$. The validity of the claim for $\alpha = 0$ is immediate. **QED.**

The likelihood for more general games is given by requiring that q be a product distribution and that Eq. 11 hold simultaneously for all players i (other than Nature players). Combining, our posterior for this more general case is

$$P(q \mid \mathscr{I}) \propto e^{\alpha S(q)} I(q \in \Delta_{\mathcal{X}}) \prod_{i=1}^{N} I(q_i \cdot U_{q_{-i}} = K(U_{q_{-i}}, b_i)). \tag{13}$$

Note that any QRE with logit exponents $\{b_i\}$ has its likelihood equal to 1.

Unlike motivations of the QRE, to motivate our choice of ϵ_i we do *not* say that each q_i must be a logit distribution. The probability density over possible q_i is not assumed to be a delta function about a logit q_i. Rather we make the weaker presumption that QRE distribution has non-zero likelihood in the single-player inference problem. That presumption motivates a $\boldsymbol{b}$-parameterized likelihood that can then be applied in the multi-player scenario. (An even weaker assumption — beyond the scope of this paper — would have $\boldsymbol{b}$ be a random variable that is sampled before the game is played.)

Define $\mathscr{I}_{\boldsymbol{b}}$ as the set of q such that $\forall i, q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$, where it is implicitly assumed that $\boldsymbol{b} \succeq \mathbf{0}$. For any such $\boldsymbol{b}$ there is always at least one $q \in \mathscr{I}_{\boldsymbol{b}}$; every QRE for the set of logit exponents $\boldsymbol{b}$ is a member of $\mathscr{I}_{\boldsymbol{b}}$. Since the support of the entropic prior is all $\Delta_{\mathcal{X}}$, this means that for any $\boldsymbol{b} \succeq \mathbf{0}$, the posterior conditioned on $q \in \mathscr{I}_{\boldsymbol{b}}$ is always non-zero at every $q^*_{\boldsymbol{b}}$. Accordingly, the posterior is well-defined, and therefore so are its local peaks, and in particular the MAP.

On the other hand, for any finite $\boldsymbol{b}$, in general there are uncountably many q's that also satisfy $\mathscr{I}_{\boldsymbol{b}}$ in addition to the QRE's. In fact, in general $P(q \mid \mathscr{I}_{\boldsymbol{b}})$ is non-zero for q's that are not products of logit distributions.

As a result of our posterior, even though the moves $\{x_i\}$ of the players are independent for any particular q (since q is a product distribution), our (!) lack of knowledge concerning the set of all the instances might result in a posterior $P(q \mid \mathscr{I})$ in which the distributions $\{q_i\}$ are statistically coupled. (Recall that q reflects the players, and P reflects our inference concerning them.) Now for the entropic prior $P(q)$ there is no statistical coupling between x_i and x_j in the prior distribution $P(x)$ (cf. Eq. 6). However the potential coupling between the $\{q_i\}$ means that in the posterior distribution, the moves are *not* statistically independent (assuming one doesn't condition on q). In such a situation, *to us*, x_i and x_j are statistically coupled. This means that in some situations the joint mixed strategy $P(x \mid \mathscr{I})$ cannot equal a Nash equilibrium of the underlying game; a Nash equilibrium is impossible.

These conclusions about the impossibility of Nash equilibria do not depend on our choice of ϵ_i, or even on our encapsulating $\mathscr{I}$ in terms of ϵ_i's. (N.b., we're explicitly allowing the case where $P(q \mid \mathscr{I})$ is restricted to Nash equilibria.) Rather they come from the fact that our prior allows non-zero probability for all of the Nash equilibria.

Example 2: Consider a common payoff symmetric game involving two players, each with move space $\{A, B\}$. Let the shared utility function be $u(A, A) = 2, u(A, B) = u(B, A) = 0, u(B, B) = 1$. This game has three Nash equilibria: $(A, A), (B, B)$, and the mixed strategy where each player makes move A with probability $1/3$. The first two of those q have entropy 0 (they are delta functions). The associated value of the entropic prior, $exp(\alpha S(q))/Z(\alpha)$, is just $[Z(\alpha)]^{-1}$. The last Nash equilibrium has entropy ln[3] - 2/3ln[2].

If we define $w(\alpha) \triangleq exp(\alpha\{\ln[3] - 2/3\ln[2]\})$, then the prior probability of the first two (pure strategy) Nash equilibria are $1/[2 + w(\alpha)]$, and the prior probability of the last (mixed strategy) Nash equilibrium is $w(\alpha)/[2 + w(\alpha)]$. Since all three equilibria have the same likelihood (namely, 1), these prior probabilities are also their respective posterior probabilities, $P(q \mid \mathscr{I})$, i.e., they are the three values of a^j. Accordingly,

$$\begin{aligned}
P(x = (A, A) \mid \mathscr{I}) &= \frac{1}{2 + w(\alpha)} + \frac{w}{(2 + w)}[\frac{1}{3}]^2 = \frac{9 + w(\alpha)}{9(2 + w(\alpha))}, \\
P(x = (B, B) \mid \mathscr{I}) &= \frac{1}{2 + w(\alpha)} + \frac{w}{(2 + w)}[\frac{2}{3}]^2 = \frac{9 + 4w(\alpha)}{9(2 + w(\alpha))}, \\
P(x = (A, B) \mid \mathscr{I}) &= P(x = (B, A) \mid \mathscr{I}) = \frac{2w}{9(2 + w(\alpha))} \qquad (14)
\end{aligned}$$

Not only is this distribution $P(x \mid \mathscr{I})$ not a Nash equilibrium; neither player i plays best-response to $P(x_{-i} \mid \mathscr{I})$. In fact, $P(x \mid \mathscr{I})$ is not even a product distribution.

4.2 The posterior q covers all Nash equilibria

Let q be a Nash equilibrium where for some player i, $R \equiv \text{supp}[q_i]$ includes multiple $x_i \in X_i$ and q_i is not uniform over R. Since q is a Nash equilibrium, $\mathbb{E}_{q_{-i}}(u^i \mid x_i)$ is uniform over R. This means that any logit distribution $\mathcal{L}_{\mathbb{E}_{q_{-i}}(u^i|.),c}(x_i)$ must be uniform across all $x_i \in R$. Since by hypothesis q_i is not uniform over R, this means that q_i cannot be described by a logit distribution. So such a Nash equilibrium q is not a QRE for any vector of rationalities $\boldsymbol{b}$, even one including infinite components. This complicates consideration of Nash equilibria in terms of QRE's, leading to the analysis of limits of QRE's as $\boldsymbol{b} \to \infty$.

Such complications do not arise in PGT, as illustrated in the following two results:

Proposition 2. *For any $q \in \Delta_{\mathcal{X}}$ there is one and only one $\boldsymbol{b}$ such that $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}$ $\forall i$. Define $B : \Delta_{\mathcal{X}} \to \mathbb{R}^N$ as that function taking any q to the associated vector of rationalities. Then B_i is differentiable everywhere in $\Delta_{\mathcal{X}}$ that it is finite.*

Proof. Pick any player i. If $q^i \cdot U^i_{q_{-i}} = \max[U^i_{q_{-i}}(x_i)]$ then $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}$ $\forall i$ iff $b_i = \infty$. Similarly $K(U^i_{q_{-i}}, b_i) = \min[U^i_{q_{-i}}(x_i)]$ iff $b_i = -\infty$. Now consider the remaining cases, where $q^i \cdot U^i_{q_{-i}} \in (\min[U^i_{q_{-i}}(x_i)], \max[U^i_{q_{-i}}(x_i)])$. Due to the bijectivity of $K(U^i_{q_{-i}}, .)$ with that codomain, we again see that there is a unique b_i such that $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}$ $\forall i$. This completes the first claim.

To establish the second claim evaluate the derivative of B_i and show that it is finite. We do this by applying the chain rule to $K(U^i_{q_{-i}}, B_i(q)) - q^i \cdot U^i_{q_{-i}} = 0$. The result for the components $q_i(x_i)$ and $q_{-i}(x_{-i})$ of the argument list of B_i are

$$\frac{\partial B_i(q_i, q_{-i})}{\partial q_i(x_i)} = \frac{U^i_{q_{-i}}(x_i)}{\frac{\partial K(U^i_{q_{-i}}, B_i)}{\partial B_i}}$$
$$\frac{\partial B_i(q_i, q_{-i})}{\partial q_{-i}(x_{-i})} = \frac{\left(q_i(x_i) \cdot \frac{\partial U^i_{q_{-i}}(x_i)}{\partial q_{-i}(x_{-i})}\right) - \frac{\partial K(U^i_{q_{-i}}, B_i)}{\partial q_{-i}(x_i)}}{\frac{\partial K(U^i_{q_{-i}}, B_i)}{\partial B_i}}$$

where the shared denominator is non-zero since B_i is finite by hypothesis and since $K(U^i_{q_{-i}}, .)$ is an increasing function of its second argument. **QED.**

We can use Prop. 2 to establish the following result:

Proposition 3. *For any $q \in \Delta_{\mathcal{X}}$ there is one and only one $\boldsymbol{b}$ such that $P(q \mid \mathscr{I}_{\boldsymbol{b}}) \neq 0$. For that $\boldsymbol{b}$, for all $q' \in \Delta_{\mathcal{X}}$,*

$$\frac{P(q \mid \mathscr{I}_{\boldsymbol{b}})}{P(q' \mid \mathscr{I}_{\boldsymbol{b}})} \geq |X|^{-\alpha} \tag{15}$$

where α is the exponent of the entropic prior.

Proof. Prop. 2 means that for every $q \in \Delta_{\mathcal{X}}$, there is one (and only one) $\boldsymbol{b}$ such that the likelihood $P(\mathscr{I}_{\boldsymbol{b}} \mid q)$ is non-zero. Since the entropic prior is non-zero for all q, this means that every q has non-zero posterior $P(q \mid \mathscr{I}_{\boldsymbol{b}})$ under exactly one $\boldsymbol{b}$, as claimed.

Given q, define $\boldsymbol{b}^* \equiv B(q)$, so that $P(\mathscr{I}_{\boldsymbol{b}^*} \mid q) = 1$. Now $P(\mathscr{I}_{\boldsymbol{b}^*} \mid q') \leq 1$ for any q'. Accordingly, the ratio in the proposition is bounded above by the ratio of the exponential prior at q to that at q'. However the ratio of $e^{\alpha S(q'')}$ between any two points q'' is bounded below by $\frac{exp(\alpha \cdot 0)}{exp(\alpha \ln(|X|))}$. **QED.**

In particular, this result holds for Nash equilibrium q; such equilibria arise for $\boldsymbol{b} = \infty$. The relative probabilities of those Nash q are given by the ratios of the associated prior probabilities. Prop. 3 also holds for any particular q infinitesimally close to one of the Nash equilibria. In this sense, the posterior probability is arbitrarily tightly restricted to any one of the Nash equilibria for some appropriate $\boldsymbol{b}$.

The picture that emerges then is that $\forall \boldsymbol{b}$, $\exists$ noni-empty proper submanifold of $\Delta_{\mathcal{X}}$ that is the support of the associated posterior. There is no overlap between those submanifolds (one for each $\boldsymbol{b}$), and their union is all of $\Delta_{\mathcal{X}}$, including the Nash equilibria q's (for which $\boldsymbol{b} = \infty$). Within any single one of the submanifolds no q has too small a posterior (cf. Prop. 3). This is because all q within a single submanifold have the same value (namely 1) of their likelihoods. Accordingly, the ratios of the posteriors of the q's within the submanifold is given by the ratios of (the exponentials of) the entropies of those q's.

4.3 The MAP q

Naively, one might presume that a QRE is the MAP of our posterior. After all, this is the case when a single player plays against Nature. Furthermore, when there are multiple players, every QRE q obeys our constraints that $\mathbb{E}_q(u^i) = \epsilon_i(U^i_q)\ \forall i$, and it maximizes the entropy of each player's strategy considered in isolation of the others. However in general a QREwill not maximize the entropy of the joint mixed strategy subject to our constraints when there are multiple players. In other words, while MAP for each individual player's strategy, in general it is not MAP for the joint strategy of all the players. The reason is that setting each separate q_i to maximize the associated entropy (subject to having q obey our invariant), in a sequence, one after the other, will not in general result in a q that maximizes the sum of those entropies. So it will not in general result in a q that maximizes the entropy of the joint system.

Proceeding more carefully, call a local maximum of the posterior that is in the interior of $\Delta_{\mathcal{X}}$ a "local peak" of the posterior. As shorthand, we introduce the following notation:

Definition 1. *For all* $j, \boldsymbol{b}, \phi \in \mathbb{R}^N$,

$$
\begin{aligned}
q_j^\dagger(x_j) &\triangleq \mathcal{L}_{U^j_{q_{-j}}, b_j}(x_j), \\
r_j(q, x_i) &\triangleq \sum_{x_j} q_j^\dagger(x_j) \mathbb{E}_q(u^j \mid x_i, x_j) \; [1 \; + \; b_j \{\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j} \times q_j^\dagger}(u^j)\}], \\
s_i(\phi, x_i) &\triangleq \sum_{j \neq i} \phi_j \left[\mathbb{E}_{q_{-i}}(u^j \mid x_i) - r_j(q, x_i)\right]
\end{aligned}
$$

Then we have the following lemma:

Lemma 1. *For a given* $\boldsymbol{b}$*, any local peak of the posterior is given by the* q_i *members of a set of values* $\{q_i \in \Delta_{X_i}, \lambda_i \in \mathbb{R}\}$ *that simultaneously solves the following equations for all* i*:*

$$
\begin{aligned}
q_i(x_i) \; &\propto \; e^{\lambda_i U^i_{q_{-i}}(x_i) + s_i(\lambda, x_i)}, \\
\mathbb{E}_q(u^i) &= K(U^i_{q_{-i}}, b_i).
\end{aligned}
$$

Proof. By examination of the posterior, its maxima are q's in $\Delta_{\mathcal{X}}$ that maximize $S(q)$ subject to the constraints in Eq. 11. (Recall that there always exist $q \in \Delta_{\mathcal{X}}$ obeying those constraints.) So the local peaks of the posterior are the critical points of the Lagrangian $\mathscr{L}(q, \{\lambda_i\}) \; = \; S(q) + \sum_i \lambda_i (q_i \cdot U^i - \epsilon_i(U^i)) + \sum_i \gamma_i (\sum_{x_i} q_i(x_i) - 1)$ that obey $q_i(x_i) > 0 \; \forall i, x_i$, where the λ_i are Lagrange parameters enforcing the constraints in Eq. 11 and the γ_i are Lagrange parameters forcing each q_i to be normalized. At any such critical point, $\forall \; i, x_i \in X_i$,

$$
\begin{aligned}
0 = \frac{\partial \mathscr{L}}{\partial q_i(x_i)} \; &= \; -1 - \gamma_i - \ln[q_i(x_i)] + \lambda_i \mathbb{E}(u^i \mid x_i) + \sum_{j \neq i} \lambda_j [\mathbb{E}(u^j \mid x_i) - \frac{\partial \epsilon_j(U^j)}{\partial q_i(x_i)}] \\
&= -1 - \ln[q_i(x_i)] + \lambda_i \mathbb{E}(u^i \mid x_i) + \\
&\quad \sum_{j \neq i} \lambda_j \left[\mathbb{E}_{q_{-i}}(u^j \mid x_i) - \sum_{x_j} \frac{\partial \epsilon_j(U^j)}{\partial U^j(x_j)} \mathbb{E}_{q_{-i,-j}}(u^j \mid x_i, x_j)\right].
\end{aligned}
$$

Accordingly, at such q's, for all players i,

$$
q_i(x_i) \propto e^{\lambda_i \mathbb{E}_{q_{-i}}(u^i | x_i) + \sum_{j \neq i} \lambda_j \left[\mathbb{E}_{q_{-i}}(u^j | x_i) - \sum_{x_j} \frac{\partial \epsilon_j(U^j)}{\partial U^j(x_j)} \mathbb{E}_{q_{-i,-j}}(u^j | x_i, x_j)\right]}
$$

where the proportionality constant enforces normalization. By inspection, for any real-valued Lagrange parameters, each such q_i does obey $q_i(x_i) > 0 \; \forall x_i$, as required.

To proceed plug in Eq. 11 and then Eq. 10 to evaluate $\frac{\partial \epsilon_j(U^j)}{\partial U^j(x_j)}$. Then interchange the order of the two differentiations, to differentiate with respect to $U^j(x_j)$ before differentiating with respect to b_j The result is

$$\frac{\partial \epsilon_j(U^j_{q_{-j}})}{\partial U^j_{q_{-j}}(x_j)} = q^\dagger_j(x_j)[1 \ + \ b_j\{U^j_{q_{-j}}(x_j) - \mathbb{E}_{q^\dagger_j}(U^j_{q_{-j}})\}]$$

where we have made explicit the dependence of each U^j on q_{-j}. Next use the definition of $U^j_{q_{-j}}$ and the fact that q is a product distribution to expand this result as

$$\frac{\partial \epsilon_j(U^j_{q_{-j}})}{\partial U^j_{q_{-j}}(x_j)} = q^\dagger_j(x_j)[1 \ + \ b_j\{\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j} \times \mathcal{L}_{U^j_{q_{-j}}, b_j}}(u^j)\}].$$

Now plug this result into the outer summands in our equation above for each $q_i(x_i)$, getting

$$\sum_{x_j} \frac{\partial \epsilon_j(U^j_{q_{-j}})}{\partial U^j_{q_{-j}}(x_j)} \mathbb{E}_q(u^j \mid x_i, x_j)$$
$$= \sum_{x_j} q^\dagger_j(x_j) \mathbb{E}_q(u^j \mid x_i, x_j) \ [1 \ + \ b_j\{\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j} \times q^\dagger_j}(u^j)\}].$$

Plugging in the definition of r_j completes the proof. **QED.**

In particular, the MAP is a local peak of the posterior. Therefore if the MAP is interior to $\Delta_{\mathcal{X}}$ it must solve the coupled set of equations given in Lemma 1.

4.4 The modes of $P(q \mid \mathscr{I})$ and the QRE's

It is illuminating to compare the conditions of Lemma 1 for q to be a local peak of the posterior to conditions for it to be a QRE: any QRE is given by the q_i members of a set of values $\{q_i \in \Delta_{X_i}, \lambda'_i \in \mathbb{R}\}$ that simultaneously solves the following equations for all i:

$$\begin{aligned} q_i(x_i) \ &\propto \ e^{\lambda'_i U^i_{q_{-i}}(x_i)}, \\ \mathbb{E}_q(u^i) &= K(U^i_{q_{-i}}, b_i) \end{aligned} \tag{16}$$

where the second equation forces $\lambda'_i = b_i \ \forall i$.[8]

This comparison suggests that in some circumstances the QRE is an approximation of the local peaks of the posterior $P(q \mid \mathscr{I})$. To confirm this, first note that, ultimately, the only free parameter in our solution for the local peak q's is $\boldsymbol{b}$. In addition, any QRE q^* is a solution to a set of coupled nonlinear equations parameterized by $\boldsymbol{b}$. In general there is a very complicated relation between the the local peak q's and the q^*'s, one that varies with $\boldsymbol{b}$ (as well as with the $\{u^j\}$, of course).

[8] To see this, use the first equation to write $\mathbb{E}_q(u^i) = K(U^i_{q_{-i}}, \lambda'_i)$, and recall that $K(.,.)$ is monotonically increasing in its second argument.

Intuitively, the reason for the difference between the two solutions is that each player i does not operate in a fixed environment, but rather in one containing intelligent players trying to adapt their moves to take into account i's moves. This is embodied in the likelihood of Eq. 11. In contrast to that likelihood, the likelihoods of the QRE each implicitly assume that the associated player i operates in a fixed environment.

Formally, the difference arises because each q_i not only appears in the first term in the argument of $I(q_i \cdot U_{q_{-i}} = K(U_{q_{-i}}, b_i))$ (which is the case in the game against Nature). It also occurs in the second arguments of $I(q_j \cdot U_{q_{-j}} = K(U_{q_{-j}}, b_j))$ for the players $j \neq i$. This means that if we change q_i, then the likelihood of Eq. 11 induces a change to q_{-i}, to have the invariant for the players other than i still be satisfied. This change to q_{-i} then induces a "second order" change to q_i, to satisfy the invariant for player i.

This second-order effect will not arise in a game against Nature, which treats the other players as fixed. This reflects the fact that such a game against Nature is an instance of decision theory, lacking the common knowledge aspect of games with multiple conflicting players.

Now in general it is not the case that for every i, $q_i(x_i)$ equals $q_i^\dagger(x_i)$ on an x_i-by-x_i basis. Indeed, if this were the case then q would be a QRE. However as an approximation we can impose the weaker condition that the differences between those distributions approximately cancel out inside the appropriate sum from Lemma 1:

$$\begin{aligned}
\mathbb{E}_{q_j^\dagger \times q_{-j,-i}}(u^j \mid x_i) &= \sum_{x_j} q_j^\dagger(x_j)\mathbb{E}_{q_{-j,-i}}(u^j \mid x_i, x_j) \\
&\approxeq \sum_{x_j} q_j(x_j)\mathbb{E}_{q_{-j,-i}}(u^j \mid x_i, x_j) \\
&= \mathbb{E}_{q_{-i}}(u^j \mid x_i).
\end{aligned} \tag{17}$$

(In particular, any QRE obeys this approximation exactly.) Making this approximation inside all r_j, for any $\phi \in \mathbb{R}^N$,

$$s_i(\phi, x_i) = -\sum_{j \neq i} \phi_j b_j \sum_{x_j} q_j^\dagger(x_j)\mathbb{E}_{q_{-j,-i}}(u^j \mid x_i, x_j)\left[\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j} \times q_j^\dagger}(u^j)\right] \tag{18}$$

which we can write as

$$s_i(\phi, x_i) = -\sum_{j \neq i} \phi_j b_j \mathrm{Cov}_{q_j^\dagger(x_j)}\left[\mathbb{E}_q(u^j \mid x_i, x_j),\ \mathbb{E}_q(u^j \mid x_j)\right]. \tag{19}$$

Combined with Lemma 1 this provides the following result:

Theorem 1. *Let q be a joint mixed strategy where $\exists\ \mu \in \mathbb{R}^N$ and $t \in \mathbb{R}$ such that simultaneously $\forall i, x_i$,*

$$\sum_{j \neq i} \mu_j b_j \, Cov_{q_j(x_j)}[\mathbb{E}_q(u^j \mid x_j, x_i), \mathbb{E}_q(u^j \mid x_j)] \;=\; (\mu_i - b_i)\mathbb{E}_q(u^i \mid x_i) + t.$$

Then the following two conditions are equivalent:
i) q is a QRE.
ii) q is a local peak of the posterior and obeys Eq. 17 exactly.

Proof. It is immediate that if q is a QRE then it obeys Eq. 17 exactly. This means that $s_i(\mu, x_i)$ equals the expression in Eq. 19 for $\phi = \mu$. Accordingly, the condition in the theorem involving a sum of covariances means that the exponent in Lemma 1 reduces to $(b_i+\lambda_i-\mu_i)\mathbb{E}_{q_{-i}}(u^i \mid x_i)-t$ for all $i, x_i \in X_i$. So by that lemma, for our q to be a local peak of the posterior it suffices for there to be a $\lambda \in \mathbb{R}^N$ such that $\mathbb{E}_q(u^i) = K(U^i_{q_{-i}}, b_i)$ and $q_i(x_i) \propto e^{(b_i+\lambda_i-\mu_i)\mathbb{E}_{q_{-i}}(u^i|x_i)}$ for all i, x_i. Since q is a QRE with exponent b_i, both of these conditions are met for $\lambda = \mu$. Therefore q is a local peak of the posterior, as claimed.

To prove the converse, plug the condition in the theorem involving a sum of covariances with $\phi = \mu$ into the expression in Eq. 19 for $s_i(\phi, x_i)$. Identifying $\lambda' = \lambda - \mu + b$, this reduces Lemma 1 to the conditions in Eq.'s 16 sufficient for q to be a QRE. **QED.**

In particular, say that $\sum_{j \neq i}(b_j)^2 \mathrm{Cov}_{q^*_j(x_j)}[\mathbb{E}_{q^*}(u^j \mid x_j, x_i), \mathbb{E}_{q^*}(u^j \mid x_j)]$ is independent of x_i $\forall i$ at some QRE q^*. Then the condition in Thm. 1 holds, with $\mu = \boldsymbol{b}$. So any such QRE is a local peak of the posterior.

Particularly for very large systems (e.g., a human economy), it may be that at some QRE q^*, $\mathbb{E}_{q^*}(u^j \mid x_j, x_i) = \mathbb{E}_{q^*}(u^j \mid x_j)$ for almost any i, j and associated moves x_i, x_j. In this situation, at the QRE the move of almost any player i has no effect on how the expected payoff to player j depends on j's move. If this is in fact the case for player i and all other players j, then the covariance for each j, x_i that occurs in Thm. 1 reduces to the variance of $\mathbb{E}_{q^*}(u^j \mid x_j)$ as one varies x_j according to q^*_j. By the discussion in Sec. 3.2 this variance is given by the partition function:

$$\mathrm{Var}_{q^*_j}(\mathbb{E}_{q^*}(u^j \mid x_j)) = \mathrm{Var}_{q^*_j}(U^j_{q^*}) \;=\; \frac{\partial^2 \ln(Z_{U^j_{q^*}}(b'_j))}{\partial (b'_j)^2}|_{b'_j = b_j}. \tag{20}$$

In particular, for $b_j \to \infty$ — perfectly rational behavior on the part of agent j — the variance goes to 0. So assume that Eq. refeq:approx holds to a very good approximation. Then if every player i is "decoupled" from all other players, in the limit that all players become perfectly rational the condition in Thm. 1 generically is met for $\mu = \boldsymbol{b}$. (The b_j-dependence in the covariance occurs in an exponent, and therefore generically overpowers the $(b_j)^2$ multiplicative factor.) So the QRE's approach the local peaks of the posterior in that situation.

On the other hand, if the players have bounded rationality, their variances are nonzero. In this case the expression in Thm. 1 is nonzero for each i, j, x_i.

Typically for fixed i the precise nonzero value of that variance will vary with x_i. In this case, Thm. 1 suggests that the QRE differs from the local peaks of the posterior, and in particular differ from the MAP.

There are many ways that these results can be extended. For example say a particular QRE is a local peak of the posterior for some $\boldsymbol{b}$. Then we can use a Laplace expansion to approximate the posterior in the vicinity of that QRE as a Gaussian projected onto the submanifold of joint mixed strategies that obey Eq. 16 [Robert and Casella, 2004]. Say that that QRE is close to the mean of the posterior over q's (e.g., this would be the case if that QRE is the MAP and the posterior is sharply peaked). Then our Gaussian approximation could be used to approximate the variance of any function of q under our posterior.

5 Other applications of PGT

Under the likelihood introduced above, if q_{-i} changes, then $U^i_{q_{-i}}$ changes, and therefore q_i may have to change. So this likelihood implicitly presumes the players have had some form of interaction to couple them (just as do conventional equilibrium concepts when they have multiple solutions). In [Wolpert, 2005] an different likelihood is introduced that involves no such coupling. This likelihood can be viewed as a novel formulation of common knowledge [Aumann, 1999, Aumann and Brandenburger, 1995, Fudenberg and Tirole, 1991].

Interestingly, as it arises with this likelihood, bounded rationality is identical to an information-theoretic cost of computation. In this sense, under this likelihood cost of computation is *derived* as a cause of bounded rationality. It is not simply imputed, as an explanation of experimentally observed bounded rationality.

While various models of bounded rationality have been found to have some experimental validity (e.g., QRE's), no model with a small number of parameters will ever hold *exactly*. This means that to analyze the rationality of human behavior in experimental settings we need a way to quantify the rationality of *any* mixed strategy in *any* environment. As elaborated in [Wolpert, 2005], PGT provides such a rationality measure, one that can be derived from first-principles involving the Kullbach-Leibler distance.

PGT is applicable to many domains beyond those considered in this paper. In particular, in work in progress, PGT has been used to derive power law distributions over the possible outcomes in unstructured bargaining. Those distributions have the Nash bargaining solution as their mode.

ACKNOWLEDGEMENTS: I would like to thank Julian Jamison, George Judge, Duncan Luce, Herb Gintis, and Larry Karp for helpful discussions.

References

[Aliprantis and Border, 2006] Aliprantis, C. and Border, K. C. (2006). *Infinite Dimensional Analysis*. Springer Verlag.

[Allais, 1953] Allais, M. (1953). *Econometrica*, 21:503–546.

[Anderson et al., 2002] Anderson, S. P., Goeree, J., and Holt, C. A. (2002). The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, 69(1):21–47.

[Aoki, 2004] Aoki, M. (2004). *Modeling Aggregate Behavior and Fluctuations in Economics : Stochastic Views of Interacting Agents*. Cambridge University Press.

[Aumann and Hart, 1992] Aumann, R. and Hart, S. (1992). *Handbook of Game Theory with Economic Applications*. North-Holland Press.

[Aumann, 1999] Aumann, R. J. (1999). Interactive epistemology ii: Probability. *Int. J. Game Theory*, 28:301–314.

[Aumann and Brandenburger, 1995] Aumann, R. J. and Brandenburger, A. (1995). Epistemic conditions for nash equilibrium. *Econometrica*, 63(5):1161–1180.

[Basar and Olsder, 1999] Basar, T. and Olsder, G. (1999). *Dynamic Noncooperative Game Theory*. Siam, Philadelphia, PA. Second Edition.

[Berger, 1985] Berger, J. M. (1985). *Statistical Decision theory and Bayesian Analysis*. Springer-Verlag.

[Bernardo and Smith, 2000] Bernardo, J. and Smith, A. (2000). *Bayesian Theory*. Wiley and Sons.

[Binmore, 1992] Binmore, K. (1992). *Fun and Games: A Text on Game Theory*. D. C. Heath and Company, Lexington, MA.

[Brock and Durlauf, 2001] Brock, W. A. and Durlauf, S. N. (2001). Interaction-based models. In *Handbook of Econometrics: Volume 5*, pages 3297–3380. Elsevier. Chapter 54.

[Camerer, 2003] Camerer, C. (2003). *Behavioral Game theory: experiments in strategic interaction*. Princeton University Press.

[Chen and Friedman, 1997] Chen, H. C. and Friedman, J. W. (1997). *Games and Economic Behavior*, 18:32–54.

[Cover and Thomas, 1991] Cover, T. and Thomas, J. (1991). *Elements of Information Theory*. Wiley-Interscience, New York.

[Dragulescu and Yakovenko, 2000] Dragulescu, A. and Yakovenko, V. (2000). Statistical mechanics of money. *Eur. Phys. J. B*, 17:723–729.

[Durlauf, 1999] Durlauf, S. (1999). How can statistical mechanics contribute to social science? *Proc. Natl. Acad. Sci. USA*, 96:10582–10584.

[Farmer et al.,] Farmer, J., Shubik, M., and Smith, D. E. Economics: The next physical science? SFI working paper 05-06-027.

[Fudenberg and Kreps, 1993] Fudenberg, D. and Kreps, D. (1993). Learning mixed equilibria. *Game and Economic Behavior*, 5:320–367.

[Fudenberg and Levine, 1993] Fudenberg, D. and Levine, D. K. (1993). Steady state learning and Nash equilibrium. *Econometrica*, 61(3):547–573.

[Fudenberg and Levine, 1998] Fudenberg, D. and Levine, D. K. (1998). *The Theory of Learning in Games*. MIT Press, Cambridge, MA.

[Fudenberg and Tirole, 1991] Fudenberg, D. and Tirole, J. (1991). *Game Theory*. MIT Press, Cambridge, MA.

[Goeree and Holt, 1999] Goeree, J. K. and Holt, C. A. (1999). Stochastic game theory: for playing games, not just doing theory. *Proceedings National Academy of Sciences*, 96:10564–10567.

[Gull, 1988] Gull, S. F. (1988). Bayesian inductive inference and maximum entropy. In *Maximum Entropy and Bayesian Methods*, pages 53–74. Kluwer Academic Publishers.

[Horn, 2003] Horn, K. S. V. (2003). Constructing a logic of plausible inference: a guide to cox's theorem. *International Journal of Approximate Reasoning*, 34(1):3–24.

[J. K. Goeree, 1999] J. K. Goeree, C. A. Holt, T. R. P. (1999). Quantal response equilibrium and overbidding in private-value auctions.

[Jaynes, 1957] Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106:620.

[Jaynes and Bretthorst, 2003] Jaynes, E. T. and Bretthorst, G. L. (2003). *Probability Theory : The Logic of Science*. Cambridge University Press.

[Kahneman, 2003] Kahneman, D. (2003). A psychological perspective on economics. *American Economic Review (Proceedings)*, 93:2:162–168.

[Kurzban and Houser, 2005] Kurzban, R. and Houser, D. (2005). Experiments investigating cooperative types in humans. *Proceedings of the National Academy of Sciences*, 102(5):1803–1807.

[List and Haigh, 2005] List, J. A. and Haigh, M. S. (2005). A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Sciences*, 102:945–948.

[Loredo, 1990] Loredo, T. (1990). From laplace to sn 1987a: Bayesian inference in astrophysics. In *Maximum Entropy and Bayesian Methods*, pages 81–142. Kluwer Academic Publishers.

[Luce, 1959] Luce, D. (1959). *Individual Choice Behavior*. Wesley.

[Luce and Raiffa, 1985] Luce, R. D. and Raiffa, H. (1985). *Games and Decisions*. Dover Press.

[Mackay, 2003] Mackay, D. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press.

[McKelvey and Palfrey, 1995] McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10:6–38.

[Meginniss, 1976] Meginniss, J. R. (1976). A new class of symmetric utility rules for gambles, subjective marginal probability functions, and a generalized Bayes' rule. *Proc. of the American Statisticical Association, Business and Economics Statistics Section*, pages 471 – 476.

[Paris, 1994] Paris, J. B. (1994). *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge University Press.

[Robert and Casella, 2004] Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer-Verlag, New York.

[Shamma and Arslan, 2004] Shamma, J. and Arslan, G. (2004). Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria. *IEEE Trans. on Automatic Control*, 50(3):312–327.

[Starmer, 2000] Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38:332–382.

[Strauss et al., 1994] Strauss, C., D.H. Wolpert, D., and Wolf, D. (1994). Alpha, evidence, and the entropic prior. In Mohammed-Djafari, A., editor, *Maximum Entropy and Bayesian Methods 1992*. Kluwer.

[Wolpert and Rajnarayan, 2007] Wolpert, D. and Rajnarayan, D. (2007). Parametric learning and monte carlo optimization. http://lanl.arxiv.org/abs/0704.1274.

[Wolpert, 2004a] Wolpert, D. H. (2004a). Information theory - the bridge connecting bounded rational game theory and statistical physics. In Braha, D., Minai, A., and Bar-Yam, Y., editors, *Complex Engineered Systems: Science meets technology*, pages 262–290. Springer.

[Wolpert, 2004b] Wolpert, D. H. (2004b). What Information theory says about best response, binding contracts, and Collective Intelligence. In A. Namatame, editor, *Proceedings of WEHIA04*. Springer Verlag.

[Wolpert, 2005] Wolpert, D. H. (2005). A predictive theory of games. http://lanl.arxiv.org/abs/nlin/0512015.

[Wolpert et al., 2006] Wolpert, D. H., Strauss, C. E. M., and Rajnayaran, D. (2006). Advances in distributed optimization using probability collectives. *Advances in Complex Systems*. in press.

[Zellner, 2004] Zellner, A. (2004). Some aspects of the history of bayesian information processing. *Journal of Econometrics*.